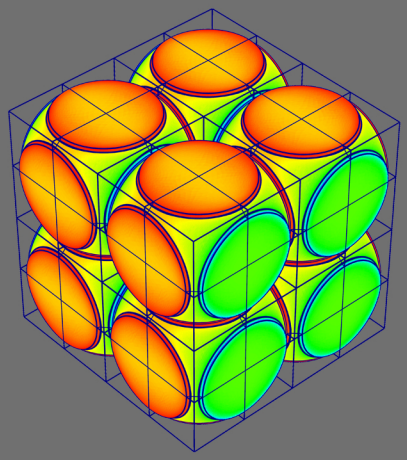
ISSN 2309-3684



# Математическое моделирование и численные методы

Силинская А.А., Богомолов А.С., Кушников В.А.
Моделирование эвакуации из помещений с учетом социальных групп и множественных выходов. Математическое моделирование и численные методы, 2025, № 3, с. 85–102.

Источник: https://mmcm.bmstu.ru/articles/384/

Параметры загрузки:
IP: 216.73.216.24
17.11.2026 03:29:22

УДК 001.891.573 DOI: 10.18698/2309-3684-2025-3-85102

# Моделирование эвакуации из помещений с учетом социальных групп и множественных выходов

© А.А. Силинская[1], А.С. Богомолов[1,2], В.А. Кушников[1,2]

[1]ФГБОУ ВО «СГУ имени Н.Г. Чернышевского», Саратов, 410012, Россия
[2]ФИЦ СНЦ РАН, Саратов, 410028, Россия

*Статья посвящена разработке многоагентной модели эвакуации, учитывающей физические характеристики агентов (возрастные категории, скорость, маневренность), уровень паники, социальные взаимодействия в группах типа «лидер–последователь» и наличие нескольких эвакуационных выходов, открывающихся с заданным интервалом (рассматривался интервал 6 секунд). Для обучения поведения агентов используется алгоритм Multi-Agent Proximal Policy Optimization (MAPPO). Используется гибридное пространство действий, сочетающее дискретный выбор выхода и непрерывное управление движением. Обучение проводится по принципу «curriculum learning» с постепенным наращиванием количества агентов. Это позволяет агентам адаптироваться к сложным сценариям с высокой скученностью и улучшить обобщающую способность модели для экспериментов с разным числом агентов. Среда представляет собой помещение заданных размеров (рассматривались помещения 15×20 м) с заданным количеством выходов определенной ширины (рассматривалось 3 выхода по 1,5 м). В модель заложена логика распространения информации о выходах. Индивидуальные агенты узнают информацию о новых открытых выходах в радиусе 5 м и передают сигнал соседям. Лидеры изначально знают обо всех доступных выходах вне зависимости от расстояния. Предусмотрен механизм распространения паники в зависимости от скученности агентов, расстояния до выхода и прошедшего времени с начала эвакуации. Введены специфические правила поведения для социальных групп: лидеры принимают стратегические решения, а пожилые последователи получают бонус к скорости при следовании за лидером. В текущей реализации выбор выхода для индивидуальных агентов основан на кратчайшем расстоянии агента до него. В социальных группах решение о выборе выхода принимается лидером на основе среднего расстояния всех агентов. Проведены вычислительные эксперименты для 40 агентов в различных сценариях: с разным числом лидеров (2–16) и без групп (индивидуальная эвакуация). Проведенные вычислительные эксперименты показали, что в рассматриваемых условиях сценарии с социальными группами приводят к более быстрой эвакуации (снижение общего времени составило около 38%). Также при групповой эвакуации наибольшее преимущество получают уязвимые агенты, в рассматриваемом случае – пожилые. Оптимальное число лидеров составляет 4–6: дальнейшее увеличение их количества не дает статистически значимых улучшений. По итогам экспериментов зафиксировано снижение количества столкновений и меньший уровень паники при таком числе лидеров. Полученные результаты демонстрируют практическую применимость подхода MAPPO к задачам анализа процессов эвакуации в реалистичных условиях.*

***Ключевые слова:*** *многоагентная модель, эвакуация, обучение с подкреплением, MAPPO, социальные группы, паника, гибридное пространство действий, лидер-последователь, возрастные категории, динамическая среда*

**Введение.** Одной из главных задач в обеспечении безопасности населения является оперативная и своевременная эвакуация людей из зданий при возникновении чрезвычайных ситуаций. Важным аспектом подготовки помещений к таким ситуациям выступает математическое моделирование процессов эвакуации в разных условиях. Оно позволяет анализировать возможные сценарии поведения людей во время аварий, определять предельно допустимое количество находящихся в помещении людей, рассчитывать время, необходимое для полной эвакуации, а также учитывать другие значимые параметры. В связи с этим вопросам моделирования эвакуации посвящено множество научных исследований и нормативных документов в различных странах мира.

Обзор публикаций показывает разнообразие подходов к моделированию эвакуации. Агентно-ориентированные модели [1-2] фокусируются на индивидуальном и групповом поведении, учитывая когнитивные и социальные аспекты: например, эмоциональные и рациональные типы агентов. Однако возможности адаптации этих моделей к изменяющимся условиям и учету физических характеристик агентов часто ограничены. Нейросетевые подходы, например, самоорганизующиеся карты Кохонена [3], предсказывают маршруты эвакуации, но не реагируют на динамические изменения среды. Геоинформационные системы [4] визуализируют оптимальные пути эвакуации с учетом инфраструктуры, но игнорируют индивидуальное поведение в условиях плотной толпы. Интеллектуальные системы оповещения [5] и прогнозирования [6] предлагают решения для управления потоками с учетом индивидуальных характеристик, но ограничиваются статическими сценариями и не моделируют сложное поведение агентов.

В [7-9] предложены многоагентные модели эвакуации с учетом физических столкновений и сложной геометрии помещений. В [8] интегрируется моделирование пожара и эвакуации с использованием модели частично упругого удара, в [9] акцентируется внимание на спонтанной эвакуации и целенаправленном поведении агентов, таких, например, как обгон. Однако в этих работах не учитываются возможные физические различия агентов, социальные взаимодействия и влияние паники на их поведение. В [10] на основании методик МЧС России учитываются возрастные группы и их физические характеристики. Модель включает механизм ожидания агентами свободного пространства при заторах, что повышает организованность эвакуации. Тем не менее, исследование также не рассматривает влияние паники и социальных связей.

Значительный прогресс в моделировании эвакуации был достигнут с применением глубокого обучения с подкреплением (DRL, Deep Reinforcement Learning). Работы [11-13] используют DRL для оптими-

зации путей эвакуации в реальном времени, представляя помещения как графы и прогнозируя заторы. Например, в [11] трансферное обучение сокращает время тренировки, но модель не учитывает столкновения и панику. В [13] агенты оптимизируют выбор выходов с учетом их пропускной способности, но представленные модели не учитывают социальные группы и физические различия. В исследовании [14] эвакуация моделируется с учетом психологических факторов, возраста и мобильности агентов. Приводится информация о точности 87% при сравнении с реальными данными. Однако данная модель не масштабируется на более сложные помещения. Многоагентное обучение с подкреплением (MARL, Multi Agent Reinforcement Learning) [15] позволяет эффективно управлять потоками в условиях конкуренции за узкие маршруты, но здесь наблюдается ощутимая потеря производительности при высокой плотности агентов. Лидер-ориентированные модели [16] на основе алгоритма оптимизации проксимальной политики (PPO, Proximal Policy Optimization) ускоряют эвакуацию за счет использования координации агентов, но упрощают при этом их поведенческие аспекты. Иерархические модели [17] обеспечивают многоуровневое управление, но требуют значительных вычислительных ресурсов. Подходы DRL и MARL представляются достаточно гибкими, но редко учитывают панику, физические столкновения и групповую динамику.

Таким образом, рассмотренные работы имеют следующие ограничения: статичность среды, игнорирование индивидуальных характеристик, столкновений и социального взаимодействия, недостаточная адаптивность к сложным помещениям с множеством выходов. Эти обстоятельства обуславливают актуальность разработки и развития новых подходов, способных масштабироваться и адаптироваться к реальным сценариям.

Задачей данной работы является разработка такой многоагентной модели, которая позволяет при моделировании эвакуации людей из помещений учитывать:

Наличие агентов трех возрастных категорий с различными физическими характеристиками (скорость, маневренность, радиус проекции).

Уровень паники и ее влияние на движение.

Социальные взаимодействия в рамках групп «лидер-последователь».

Наличие нескольких выходов, открывающихся в разное время.

В работе уделено внимание использованию гибридного пространства действий, включающего дискретный выбор действия и непрерывное движение. Работа обученной модели сравнивается в случаях одиночной эвакуации и эвакуации в социальных группах с различным числом лидеров.

**Используемые модели и алгоритмы.** Алгоритм Multi-Agent Proximal Policy Optimization (MAPPO) — это алгоритм обучения с подкреплением, адаптирующий Proximal Policy Optimization (PPO) [18] для многоагентных систем. Используется парадигма централизованного обучения с децентрализованным исполнением (CTDE, Centralized Training with Decentralized Execution), где агенты обучаются с доступом к глобальной информации, но действуют независимо, полагаясь только на локальные наблюдения [19].

В MAPPO каждый агент имеет собственную политику $\pi_{\theta_i}(a_i \mid s_i)$, где $s_i$ – локальное наблюдение, а $a_i$ – действие. Во время обучения используется общая функция ценности $V(s)$, оценивающая ожидаемую совокупную награду для глобального состояния $s$. Обновление политики основано на обрезанной суррогатной цели:

$$L^{\text{clip}}(\theta_i) = \mathbb{E}\left[ \min\left( r_t(\theta_i) A_t, \text{clip}\left( r_t(\theta_i), 1-\varepsilon, 1+\varepsilon \right) A_t \right) \right],$$

где

$$r_t\left(\theta_i\right) = \frac{\pi_{\theta_i}\left(a_t|s_t\right)}{\pi_{\theta_i^{\text{old}}}\left(a_t|s_t\right)},$$

— отношение вероятностей текущей и старой политики, $A_t$ — оценка преимущества, вычисленная с помощью Generalized Advantage Estimation (GAE), а $\varepsilon$ — коэффициент отсечения. Коэффициент отсечения определяет степень изменения политики за одно обновление: меньшие значения делают обновление более консервативным, предотвращая резкие изменения, которые могут дестабилизировать обучение.

Значение GAE ($A_t$) вычисляется как взвешенная сумма временных разностей:

$$A_t = \sum_{l=0}^{\infty} \left(\gamma\lambda\right)^l \delta_{t+l}, \quad \delta_t = r_t + \gamma V\left(s_{t+1}\right) - V\left(s_t\right),$$

где $\delta_t$ — временная разность, $r_t$ — награда, $\gamma$ — коэффициент дисконтирования, а $\lambda$ — параметр сглаживания, который балансирует между смещением и дисперсией оценки [18]. Оценка GAE позволяет более точно оценивать преимущество, учитывая долгосрочные последствия действий.

Для поддержания исследования добавляется энтропийный бонус, который поощряет разнообразие действий. Энтропия политики $H\left(\pi_{\theta_i}\right)$ добавляется к целевой функции с коэффициентом $\beta$:

$$L^{\text{entropy}}\left(\theta_i\right)=\beta E\left[H\left(\pi_{\theta_i}\left(\cdot|s_t\right)\right)\right],$$

где

$$H\left(\pi_{\theta_i}\right)=-\sum\nolimits_a \pi_{\theta_i}\left(a|s_t\right)\log\pi_{\theta_i}\left(a|s_t\right)$$

— для дискретных действий. Коэффициент $\beta$ регулирует степень исследования: большие значения усиливают случайность действий, что полезно на ранних стадиях обучения, но может замедлить сходимость на поздних этапах.

Функция потерь критика минимизирует среднеквадратичную ошибку между предсказанной ценностью $V\left(s\right)$ и фактическим возвратом $R_t$:

$$L^{\text{critic}}=E\left[\left(V\left(s_t\right)-R_t\right)^2\right],$$

где

$$R_t=\sum\nolimits_{k=t}^{T}\gamma^{k-t}r_k$$

– дисконтированный возврат. Эта функция позволяет критику оценивать ценность состояния, что улучшает стабильность обучения политики.

В [19] MAPPO был протестирован в средах кооперативных многоагентных игр, где агенты, представленные точками в двумерном пространстве, должны достигать целей, избегая столкновений, что напоминает пешеходную динамику. Было показано, что производительность с MAPPO лучше, чем с off-policy алгоритмами. Этот алгоритм более эффективен в координации действий множества агентов в условиях динамических взаимодействий.

В [20] MAPPO был использован для улучшения алгоритма ORCA (Optimal Reciprocal Collision Avoidance) для автономных мобильных роботов, работающих в толпе. Метод позволил обучать роботов принимать оптимальную скорость избегания по отношению к каждому человеку в толпе, оценивая влияние этих скоростей на итоговую траекторию. Симуляция показала, что метод показал более высокий уровень успешной навигации по сравнению с базовым ORCA, при этом сохраняя время выполнения. Результат продемонстрировал способность MAPPO успешно работать в динамических средах.

В [21] была предложена модификация алгоритма MAPPO – MAPPO-PIS (Multi-Agent Proximal Policy Optimization with Prior Intent Sharing) для управления автономными транспортными средствами. Моделировалась задача принятия решений в сложных сценариях слияния транспортных потоков с транспортными средствами различных типов с различными стилями вождения. Предложенная модель ис-

пользует модули генерации намерений и модуль повышения безопасности. Результаты показали значительное улучшение безопасности по сравнению с другими многоагентными алгоритмами обучения с подкреплением (MARL). Сделан вывод, что MAPPO подходит для обработки гетерогенных агентов с разными поведенческими характеристиками.

В [22] предлагается модификация AM-MAPPO (Action-Mask Multi-Agent Proximal Policy Optimization) для решения задачи слияния беспилотных транспортных средств на въездных рампах. Было рассмотрено использование гибридного пространства действий, где комбинируются дискретные решения (выбор полосы) и непрерывные действия (регулировка скорости). Продемонстрирована устойчивость алгоритма к нестационарности среды, а результаты превзошли другие алгоритмы MARL в сложных сценариях слияния.

Подытоживая сказанное, отметим, что алгоритм на основе MAPPO выбран для решения поставленной задачи по следующим причинам. Во-первых, он эффективно масштабируется на большое число агентов [19]. Во-вторых, CTDE позволяет учитывать глобальные взаимодействия (например, столкновения или групповую координацию) при обучении, сохраняя децентрализованное исполнение, соответствующее реальным сценариям, где агенты не имеют полного доступа к состоянию среды. В-третьих, MAPPO стабильно обучает гетерогенные политики, что подходит для агентов с различными типами (подростки, взрослые, пожилые в нашем случае) и гибридным пространством действий (дискретный выбор выхода и непрерывное движение в нашем случае). Наконец, устойчивость MAPPO к нестационарности многоагентных сред дает ему преимущество перед другими алгоритмами, такими как MADDPG [23], в задачах с динамическими взаимодействиями [19].

**Развитая математическая модель.** Предлагаемая математическая модель включает акторскую и критическую сети. Пространство действий гибридное: дискретный компонент определяет выбор выхода, непрерывный — вектор движения в двумерном пространстве. Такой подход позволяет реалистично отразить реалии эвакуации, где агенты принимают стратегические решения о цели (подходящем выходе) и тактически маневрируют в толпе. Подобные гибридные пространства действий показали свою эффективность в задачах с иерархическим управлением [24].

Акторская сеть принимает вектор наблюдений (положение, скорость, тип агента, уровень паники, информация о выходах). Она состоит из двух скрытых слоев по 256 нейронов с активацией ReLU и dropout-слоями (вероятность отключения $p = 0,2$), которые предотвращают переобучение модели при высокой размерности данных и вариативности сценариев, улучшая обобщающую способность [25-26].

Выход делится на две «головы» – дискретную и непрерывную. Дискретная выдает логиты $l_k$ для выбора выхода $k$:

$$l_k = W_d h + b_d,$$

где $h$ — выход последнего скрытого слоя, $W_d$ и $b_d$ — веса и смещение. Логиты преобразуются в вероятности через softmax – функцию, которая нормализует входные значения в вероятностное распределение, где сумма всех вероятностей равна 1:

$$\pi_{\text{disc}}\left(k|s\right) = \frac{\exp\left(l_k\right)}{\sum_j \exp\left(l_j\right)}.$$

Непрерывная «голова» выдает среднее $\mu$ и логарифм стандартного отклонения $\log\sigma$ гауссовского распределения вектора движения $\left(v_x, v_y\right)$:

$$\mu = W_c h + b_c,\ \log\sigma = W_\sigma h + b_\sigma,$$

где $W_c, b_c, W_\sigma, b_\sigma$ — обучаемые параметры.

Действие $a_{\text{cont}} = \left(v_x, v_y\right)$ сэмплируется из распределения $N(\mu,\sigma)$, причем $\log\sigma$ ограничено диапазоном $\left[-2,0;2,0\right]$ для контроля дисперсии.

Сеть критика оценивает ценность состояния, имеет аналогичную архитектуру без dropout-слоев и один выходной нейрон для $V\left(s\right)$. Оптимизация проводится через Adam – адаптивный алгоритм градиентного спуска, который использует моментные оценки первого и второго порядка для ускорения сходимости и стабильного обновления весов. Скорость обучения актора $\alpha_{\text{actor}} = 5\times10^{-5}$, критика $\alpha_{\text{critic}} = 1\times10^{-3}$. Более высокая скорость обучения критика позволяет ему быстрее адаптироваться к изменениям наград, что улучшает стабильность обучения политики [18].

*Обучение модели.* Обучение модели происходит как серия 50 эпизодов, симулирующих эвакуацию до ее завершения или достижения 5000 шагов. Применяется принцип curriculum learning [27]: число агентов начинает линейно увеличиваться с 20 до 50 начиная с 25 эпизода. Это позволяет сначала освоить простые сценарии с малым числом взаимодействий, где основная цель агента – идти к выходу, а затем масштабироваться до более сложных сценариев с высокой скученностью агентов и более тонкими взаимодействиями агентов. Такой подход ускоряет обучение и повышает обобщающую способность модели [27]. Предусмотрено несколько типов агентов, которые разли-

чаются скоростью, маневренностью и радиусом проекции (таблица 1). В эпизодах симуляции используется фиксированная пропорция возрастного состава агентов: 30% подростков, 50% взрослых, 20% пожилых.

*Таблица 1*

**Характеристики агентов по их типам**

| Тип агента | Скорость (м/с) | Маневренность | Радиус проекции (м) |
|---|---|---|---|
| Подросток | 1,5 | 1,0 | 0,2 |
| Взрослый | 1,2 | 0,8 | 0,23 |
| Пожилой | 0,8 | 0,7 | 0,23 |

В случае с социальными группами число лидеров в каждом эпизоде варьируется случайно от двух до половины количества агентов. При этом лидерами могут быть только взрослые. Случайный выбор числа лидеров в каждом эпизоде позволяет повысить обобщающую способность модели.

Траектории сохраняются в буфер для MAPPO, обновление сетей проводится с мини-батчами по 1024 перехода за 20 эпох. Функция потерь включает обрезанную суррогатную цель, энтропийный бонус с коэффициентом $\beta = 0,02$, и GAE с $\lambda = 0,95$. Параметр $\lambda$ в GAE определяет баланс между краткосрочными и долгосрочными оценками преимущества: значение 0,95 делает оценку более устойчивой к шуму в наградах [18]. Градиенты критика обрезаются до нормы 0,5, а возвраты нормализуются по стандартному отклонению для стабилизации обучения при вариативных наградах [18]. Стандартное отклонение непрерывных действий уменьшается экспоненциально, обеспечивая переход от активного исследования к точным движениям, что поддерживается практиками управления исследованием в RL [28-29]. Валидация проводится каждые 5 эпизодов, оценивая время эвакуации, панику и столкновения.

*Правила поведения агентов.* В качестве среды выступает помещение $15 \times 20$ м с тремя выходами шириной 1,5, один из выходов открыт изначально, а остальные открываются с интервалом 6 сек. Полученная площадь 300 м$^2$ характерна для различных общественных помещений, таких как конференц-залы, офисные помещения, учебные аудитории, рестораны или торговые залы. Согласно [30], для помещений с одновременным пребыванием более 50 человек требуется не менее двух выходов, а минимальная ширина эвакуационного выхода при пребывании в помещении более 15 человек должна быть не меньше 1,2 м. Агенты размещаются в среде случайно, при этом в сценариях с социальными группами лидеры располагаются первыми, а последователи – вокруг них, образуя кластеры (рис. 1).

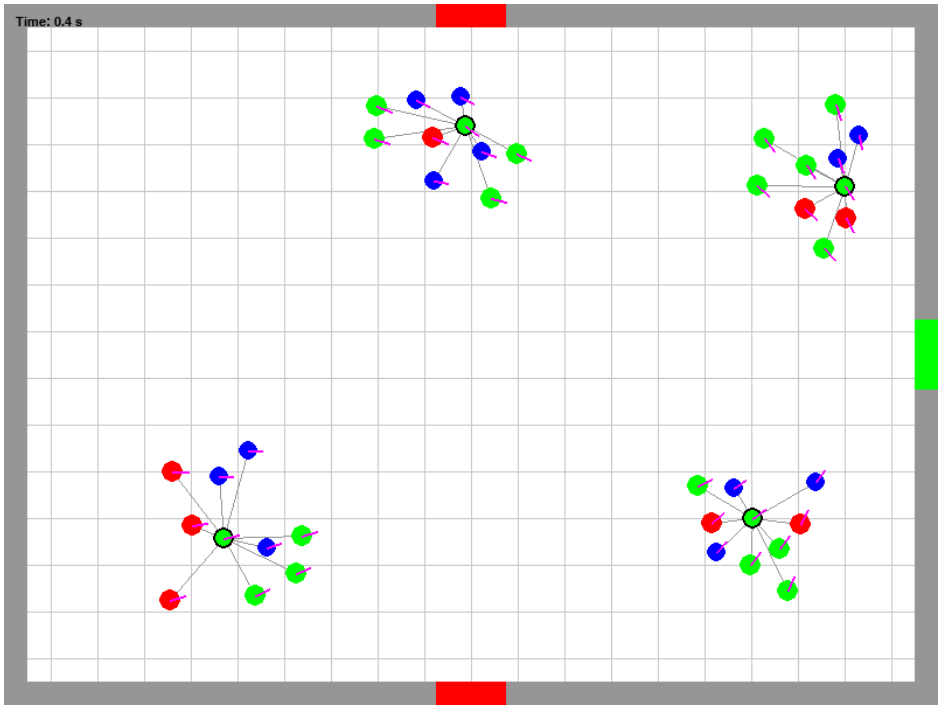



**Рис. 1.** Визуализация среды с 40 агентами и 4 лидерами

Модель предусматривает влияние паники на маневренность агентов, ее рост зависит от близости других агентов и расстояния до выхода [31], при этом в социальных группах паника растет медленнее [32]. Высокий уровень паники снижает маневренность агентов [33].

Важной особенностью логики социальных групп в модели является получение пожилыми агентами бонуса к скорости 50% при следовании за лидером. Такая особенность связана с предположением о способности лидера управлять скоростью агентов, в частности помогать уязвимым участникам движения. Такое предположение подтверждается экспериментами, проведенными в [32].

Агенты узнают о выходах в радиусе 5 м, лидеры — независимо от расстояния, все агенты передают информацию соседям в радиусе 2 м. В группах лидер выбирает выход по среднему расстоянию для последователей, но агенты (кроме пожилых) могут выбрать ближе расположенный выход. Одиночные агенты выбирают ближайший выход. У выходов (радиус 3 м) активируется очередь: агенты сортируются по расстоянию, скорость корректируется, желаемое расстояние — 0,75 м для пожилых, 0,5 м для остальных.

*Система наград*. Система наград в модели состоит из следующих компонент:

1. базовая награда: небольшой штраф для стимулирования быстрой эвакуации;

2. награда за эвакуацию: большое поощрение агентов за успешную эвакуацию с учетом скорости эвакуации (от 20 до 30);

3. награда за движение: небольшой штраф за резкое изменение направления или неподвижность;

4. награда за близость к выходу: поощрение движения к выходу, пропорционально размеру комнаты;

5. награда за смену выхода: поощрение агентов за выбор более близкого или менее переполненного выхода;

6. групповая награда: поощрение координации в группу, мотивирующее лидеров вести последователей, а последователей – следовать за лидером;

7. награда за панику: при высоком уровне паники, но близкому расположению к выходу, начисляется штраф; в остальных случаях (низкий уровень паники и большое расстояние от выхода) начисляется награда;

8. награда за скученность: штраф за нахождение в переполненных зонах или столкновения, особенно близи выходов;

9. награда за застревание: штраф за длительное отсутствие движения, особенно вблизи выхода;

10. коллективная награда: большое поощрение за успешное завершение эвакуации всеми агентами.

**Вычислительные эксперименты**. Для обучения модели и проведения вычислительных экспериментов было разработано программное обеспечение на языке Python. Использовались библиотеки: PyTorch для реализации алгоритма MAPPO, NumPy для работы с числовыми данными, Pygame для визуализации симуляции. После обучения модель запускалась для нескольких сценариев: без групп и с группами, где число лидеров варьировалось от 2 до 16 с шагом 2. Симуляции проводились для 40 агентов, для каждого сценария запущено по 50 эпизодов, для каждого эпизода измерялось среднее и общее время эвакуации для всех агентов и отдельно для каждого типа, затем значения усреднялись.

Результаты вычислительных экспериментов представлены на рис. 2.

Как показано на рис. 2, в целом в рамках рассматриваемых условий эвакуация с социальными группами проходит значительно быстрее, чем эвакуация без групп. При этом эвакуация с двумя лидерами в среднем потребовала времени меньше, чем эвакуация при любом другом сценарии.

Как показано на рис. 3, наименьшее преимущество от эвакуации в социальных группах получили самые быстрые агенты — подростки, наибольшее — пожилые.

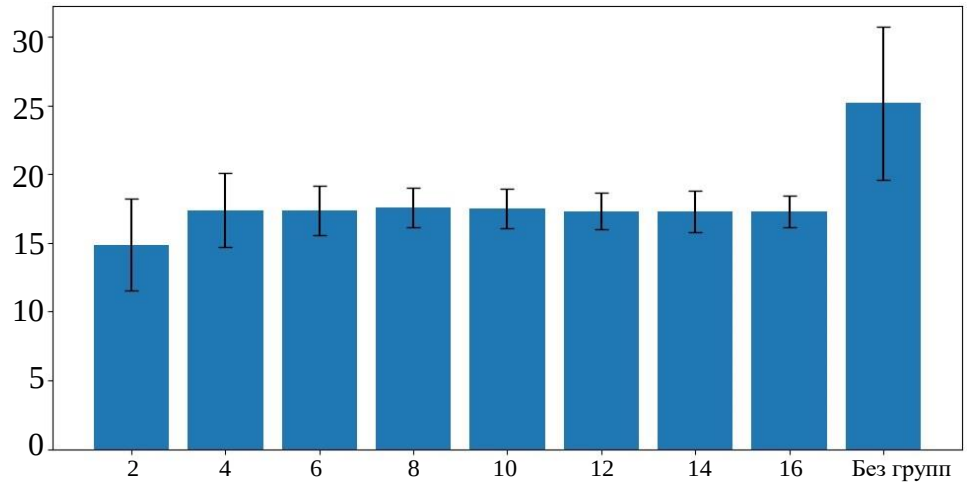



**Рис. 2.** Время эвакуации в сценариях с различным количеством лидеров. По вертикальной оси на рисунках 2 и 3 – общее время эвакуации, по горизонтальной – количество лидеров

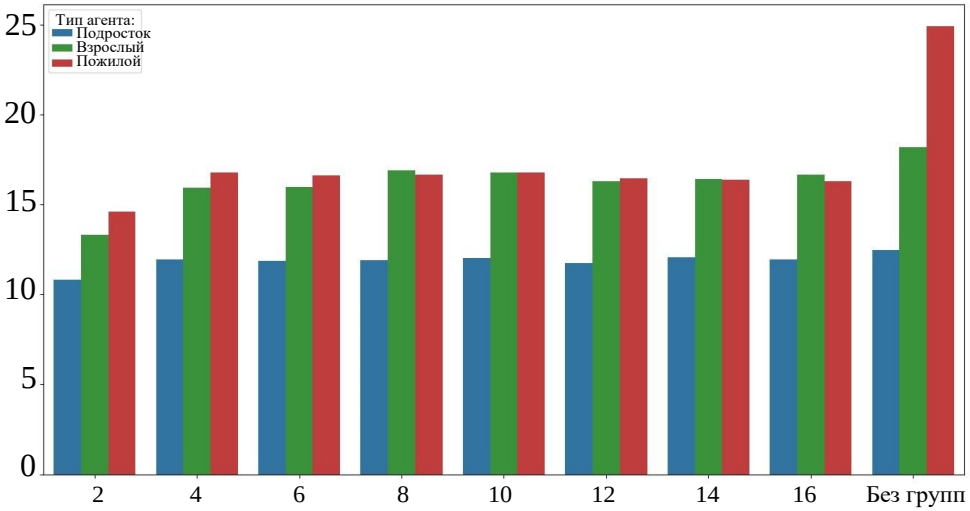



**Рис. 3.** Общее время эвакуации

При этом действительно, при эвакуации без групп пожилые агенты вносят наибольший вклад в общее время эвакуации, а в сценариях с группами полное время их эвакуации превышает время взрослых агентов незначительно.

**Обсуждение результатов**. Наибольшее среднеквадратичное отклонение полного времени эвакуации наблюдается в сценарии без групп и с двумя лидерами. В первом случае такой разброс может объясняться общей нескоординированностью агентов, а также влиянием случайного расположения агентов при инициализации среды.

Наибольшее влияние на увеличение длительности времени эвакуации оказывают пожилые агенты (рис. 3), обладающие самой низкой скоростью и маневренностью. Соответственно, при случайном расположении агентов в среде они могут появиться как на большом расстоянии от выходов, так и очень близко, что влияет на время эвакуации от эпизода к эпизоду.

Похожая ситуация с влиянием размещения агентов в среде может наблюдаться при малом количестве лидеров, в частности, при двух. Группы изначально располагаются кластерами (рис. 1), поэтому чем меньше количество лидеров или групп, тем больше вероятность возникновения ситуации, где одна из групп при инициализации среды в эпизоде расположена сильно ближе или сильно дальше от выхода.

Помимо этого, можно увидеть, что в сценариях без групп и с двумя лидерами значительно выше уровень паники. В первом случае это обуславливается общей нескоординированностью и более быстрым ростом паники у агентов без групп, во втором случае группы имеют очень большую плотность, из-за которой паника растет. По этим же причинам в данных сценариях наблюдается наибольшее число столкновений, хотя различия не настолько масштабны по сравнению с остальными сценариями. Показатели при эвакуации с 4–16 лидерами отличаются на уровне погрешности за исключением среднеквадратичного отклонения общего времени эвакуации в сценарии с 4 лидерами, что обусловлено той же причиной, что и в сценарии с двумя лидерами.

**Выводы и рекомендации.** В работе предложен способ моделирования процесса эвакуации на основе алгоритма MAPPO. Особенностью модели является возможность учитывать гетерогенность агентов, наличие нескольких выходов, открывающихся в разное время, влияние паники на движение агентов, наличие социальных групп вида «лидер-последователь». Было уделено внимание использованию гибридного пространства действий для повышения реализма модели.

Полученные результаты говорят о том, что, согласно построенной модели, эвакуация с социальными группами дает значительные преимущества во времени эвакуации, особенно для пожилых агентов. По результатам моделирования, в рассматриваемых условиях увеличивать количество лидеров свыше 4–6 становится бесполезным: метрики изменяются незначительно. При этом наиболее быстро эвакуация происходила с двумя лидерами, но стоит учитывать тот факт, что группы были организованы в кластеры при инициализации среды, в реальной ситуации агентам может потребоваться время, чтобы добраться до лидера, прежде чем приступать к движению к выходу. Эпизоды с 4–6 лидерами завершались в среднем на 2,5 секунд позже, однако имеют более стабильные показатели уровня паники, столкновений, среднеквадратичного отклонения общего и среднего времени эвакуации, чем эпизоды с двумя лидерами.

В рамках рассматриваемых условий для безопасной эвакуации всех агентов рекомендуется наличие 4–6 лидеров, помогающих своим последователям добраться до выхода. При этом наблюдается умеренный уровень паники и сравнительно небольшое число столкновений в процессе избегания скученности.

ЛИТЕРАТУРА

[1] Коткова Е.А., Матвеев А.В., Нефедьев С.А. и др. Агентное моделирование процесса эвакуации людей при пожарах в зданиях: обзор подходов и исследований. *Современные наукоемкие технологии*, 2023, № 10, с. 55–62.

[2] Zia K., Ferscha A. An agent-based model of crowd evacuation: combining individual, social and technological aspects. *Proceedings of the 2020 ACM SIGSIM conference on principles of advanced discrete simulation*. New York, Association for Computing Machinery, 2020, pp. 129–140.

[3] Коткова Е.А. Перспективы применения искусственных нейронных сетей при моделировании процесса эвакуации. *Пожарная и техносферная безопасность: проблемы и пути совершенствования*, 2020, № 1(5), с. 359–361.

[4] Суханов В.О., Кузьмин А.И., Скороходов Д.В. Геоинформационная система поддержки принятия решений на эвакуацию населения. *Пожарная безопасность: проблемы и перспективы*, 2019, т. 1, № 10, с. 411–413.

[5] Сажин И.С., Головенко Е.Л., Чаниев Б.Ю. и др. Интеллектуальная система оповещения и управления эвакуацией людей на основе информационного моделирования чрезвычайных ситуаций в здании. *Наука, техника и образование*, 2021, № 4(79), с. 40–44.

[6] Коткова Е.А., Матвеев А.В. Методика интеллектуального прогнозирования эффективности управления эвакуацией людей из общественных зданий. *Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России*, 2021, № 4, с. 107–120.

[7] Цвиркун А.Д., Резчиков А.Ф., Самарцев А.А. и др. Интегрированная модель динамики распространения опасных факторов пожара в помещениях и эвакуации из них. *Вестник компьютерных и информационных технологий*, 2019, т. 2, № 176, с. 47–54.

[8] Цвиркун А.Д., Резчиков А.Ф., Самарцев А.А. и др. Система интегрированного моделирования распространения опасных факторов пожара и эвакуации людей из помещений. *Автоматика и телемеханика*, 2022, № 5, с. 26–42.

[9] Samartsev A., Ivaschenko V., Rezchikov A., et al. Multiagent model of people evacuation from premises while emergency. *Advances in Systems Science and Applications*, 2019, vol. 19, no. 1, pp. 98–115.

[10] Гамаюнова В.О., Богомолов А.С., Кушников В.А. и др. Мультиагентное моделирование эвакуации из помещений с учетом столкновений агентов. *Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика*, 2025, т. 25, № 1, с. 106–115.

[11] Rosa A.C., Falqueiro M.C., Bonacin R., et al. EvacuAI: An Analysis of Escape Routes in Indoor Environments with the Aid of Reinforcement Learning. *Sensors*, 2023, vol. 23, no. 21, art. no. 8892.

[12] Ünal A.E., Gezer C., Pak B.K., et al. Generating emergency evacuation route directions based on crowd simulations with reinforcement learning. *2022 Innovations in Intelligent Systems and Applications Conference ASYU*. Antalya: IEEE, 2022, pp. 1–6.

[13] Xu D., Huang X., Mango J., et al. Simulating multi-exit evacuation using deep reinforcement learning. *Transactions in GIS*, 2021, vol. 25, no. 3, pp. 1542–1564.

[14] Malebary S.J., Basori A.H., Soliman alkayal E. Reinforcement learning for Pedestrian evacuation Simulation and Optimization during Pandemic and Panic situation. *Journal of Physics: Conference Series*, 2021, vol. 1817, no. 1, art. no. 012008.

[15] Komatsu H. Multi-agent reinforcement learning using echo-state network and its application to pedestrian dynamics. *arXiv preprint arXiv:2312.11834*, 2023.

[16] Sinpan N., Sasithong P., Chaudhary S., et al. Simulative Investigations of Crowd Evacuation by Incorporating Reinforcement Learning Scheme. *ICACS '22: Proceedings of the 6th International Conference on Algorithms, Computing and Systems.* New York: Association for Computing Machinery, 2022, pp. 1–5.

[17] Hassanpour S., Rassafi A.A., González V.A., et al. A hierarchical agent-based approach to simulate a dynamic decision-making process of evacuees using reinforcement learning. *Journal of choice modelling*, 2021, vol. 39, art. no. 100288.

[18] Schulman J., Wolski F., Dhariwal P., et al. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[19] Yu C., Velu A., Vinitsky E., et al. The Surprising Effectiveness of PPO in Cooperative Multi-Agent Games. *arXiv preprint arXiv:2103.01955*, 2022.

[20] Liu Z., Yao C., Na W., et al. MAPPO-Based Optimal Reciprocal Collision Avoidance for Autonomous Mobile Robots in Crowds. *2023 IEEE International Conference on Systems, Man, and Cybernetics SMC*. Honolulu: IEEE, 2023, pp. 3907–3912.

[21] Guo Y., Liu J., Yu R., et al. MAPPO-PIS: A Multi-agent Proximal Policy Optimization Method with Prior Intent Sharing for CAVs Cooperative Decision-Making. *Computer Vision – ECCV 2024 Workshops*. Cham: Springer Nature Switzerland, 2025, pp. 244–263.

[22] Shixin Z., Feng P., Anni J., et al. The unmanned vehicle on-ramp merging model based on AM-MAPPO algorithm. *Scientific Reports*, 2024, vol. 14, no. 1, art. no. 19416.

[23] Lowe R., Wu Y., Tamar A., et al. Multi-agent actor-critic for mixed cooperative-competitive environments. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook: Curran Associates Inc, 2017, pp. 6382–6393.

[24] Xiong J., Wang Q., Yang Z., et al. Parametrized Deep Q-Networks Learning: Reinforcement Learning with Discrete-Continuous Hybrid Action Space. *arXiv preprint arXiv:1810.06394*, 2018.

[25] Srivastava N., Hinton G., Krizhevsky A., et al. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 2014, vol. 15, no. 56, pp. 1929–1958.

[26] Gal Y., Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *Proceedings of The 33rd International Conference on Machine Learning*. New York: PMLR, 2016, pp. 1050–1059.

[27] Narvekar S., Peng B., Leonetti M., et al. Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey. *Journal of Machine Learning Research*, 2020, vol. 21, no. 181, pp. 1–50.

[28] Sutton R.S., Barto *A. Reinforcement learning: an introduction. Second edition.* Cambridge, The MIT Press, 2018, 552 p.

[29] Lillicrap T.P., Hunt J.J., Pritzel A., et al. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[30] Об утверждении свода правил СП 1.13130 «Системы противопожарной защиты. Эвакуационные пути и выходы»: Приказ МЧС России от 19 марта 2020 г. N 194. Москва. 2020.

[31] Trivedi A., Rao S. Agent-Based Modeling of Emergency Evacuations Considering Human Panic Behavior. *IEEE Transactions on Computational Social Systems*, 2018, vol. 5, no. 1, pp. 277–288.

[32] Ding N., Sun C. Experimental study of leader-and-follower behaviours during emergency evacuation. *Fire Safety Journal*, 2020, vol. 117, art. no. 103189.

[33] Wang L., Zheng J., Zhang X., et al. Pedestrians behavior in emergency evacuation: Modeling and simulation. *Chinese Physics B*, 2016, vol. 25, no. 11, art. no. 118901.

Статья поступила в редакцию 11.06.2025

Ссылку на эту статью просим оформлять следующим образом:

Силинская А.А., Богомолов А.С., Кушников В.А. Моделирование эвакуации из помещений с учетом социальных групп и множественных выходов. *Математическое моделирование и численные методы*, 2025, № 3, с. 85–102.

**Силинская Анна Александровна —** магистрант Саратовского национального исследовательского государственного университета имени Н.Г. Чернышевского. e-mail: asilinskaya@gmail.com

**Богомолов Алексей Сергеевич** — д-р техн. наук, доцент, профессор Саратовского национального исследовательского государственного университета имени Н.Г. Чернышевского, ведущий научный сотрудник, заведующий лабораторией ФГБУН «Саратовский научный центр Российской академии наук». e-mail: bogomolov@iptmuran.ru

**Кушников Вадим Алексеевич** — д-р техн. наук, профессор, профессор Саратовского национального исследовательского государственного университета имени Н.Г. Чернышевского, главный научный сотрудник, заведующий лабораторией ФГБУН «Саратовский научный центр Российской академии наук». e-mail: kushnikoff@iptmuran.ru

# Simulation model of information interaction in a population of agents

© N.V.Belotelov[1,2], E.A.Budanov[2]

[1]Saratov State University, Saratov, 410012, Russia
[2]FSBI FRC SSC RAS, Saratov, 410028, Russia

*The article is devoted to the development of a multi-agent evacuation model that takes into account the physical characteristics of agents (age categories, speed, maneuverability), the level of panic, social interactions in groups of the "leader-follower" type, and the presence of several evacuation exits opening at a given interval (an interval of 6 seconds was considered). The Multi-Agent Proximal Policy Optimization (MAPPO) algorithm is used to train the behavior of agents. A hybrid action space is used, combining discrete output selection and continuous motion control. Training is carried out according to the curriculum learning principle: with a gradual increase in the number of agents. This allows agents to adapt to complex scenarios with high crowding and im-*

*proves the generalization ability of the model for experiments with different numbers of agents. The environment is a room of given dimensions (rooms of 15×20 m were considered) with a given number of exits of a certain width (3 exits of 1.5 m each were considered). The model includes the logic of disseminating information about exits. Individual agents learn about new open exits within a radius of 5 m and transmit the signal to their neighbors. Leaders initially know about all available exits regardless of the distance. A mechanism is provided for spreading panic depending on the crowding of agents, the distance to the exit, and the time elapsed since the start of the evacuation. Specific rules of behavior for social groups are introduced: leaders make strategic decisions, and elderly followers receive a speed bonus when following the leader. In the current implementation, the choice of exit for individual agents is based on the shortest distance from the agent to it. In social groups, the decision to choose an exit is made by the leader based on the average distance of all agents. Computational experiments were conducted for 40 agents in various scenarios: with a different number of leaders (2–16) and without groups (individual evacuation). The computational experiments showed that under the considered conditions, scenarios with social groups lead to faster evacuation (the total time was reduced by about 38%). Also, during group evacuation, vulnerable agents receive the greatest advantage, in this case, the elderly. The optimal number of leaders is 4–6: further increase in their number does not provide statistically significant improvements. According to the results of the experiments, a decrease in the number of collisions and a lower level of panic with this number of leaders was recorded. The obtained results demonstrate the practical applicability of the MAPPO approach to the problems of analyzing evacuation processes in realistic conditions.*

***Keywords**: multi-agent model, evacuation, reinforcement learning, MAPPO, social groups, panic, hybrid action space, leader-follower, age categories, dynamic environment*

## REFERENCES

[1] Kotkova E.A., Matveev A.V., Nefedev S.A., et al. Agentnoe modelirovanie protsessa evakuatsii lyudey pri pozharakh v zdaniyakh: obzor podkhodov i issledovaniy [Agent modeling of the process of people evacuation during fire in buildings: a review of approaches and research]. *Sovremennye naukoemkie tekhnologii*, 2023, no. 10, pp. 55–62.

[2] Zia K., Ferscha A. An agent-based model of crowd evacuation: combining individual, social and technological aspects. *Proceedings of the 2020 ACM SIGSIM conference on principles of advanced discrete simulation*. New York, Association for Computing Machinery, 2020, pp. 129–140.

[3] Kotkova E.A. Perspektivy primeneniya iskusstvennykh neironnykh setey pri modelirovanii protsessa evakuatsii [Perspectives in applying artificial neural networks in modeling the evacuation process]. *Pozharnaya i tekhnosfernaya bezopasnost': problemy i puti sovershenstvovaniya*, 2020, no. 1(5), pp. 359–361.

[4] Sukhanov V.O., Kuzmin A.I., Skorokhodov D.V. Geoinformatsionnaya sistema podderzhki prinyatiya resheniy na evakuatsiyu naseleniya [Geoinformation system support decision-making on evacuation of the population]. *Pozharnaya bezopasnost': problemy i perspektivy*, 2019, vol. 1, no. 10, pp. 411–413.

[5] Sazhin I.S., Golovenko E.L., Chaniev B.Yu., et al. Intellektual'naya sistema opoveshcheniya i upravleniya evakuatsiyey lyudey na osnove informatsionnogo modelirovaniya chrezvychaynykh situatsiy v zdanii [Intelligent alerting and evacuation management system based on information emergency simulation building]. *Nauka, tekhnika i obrazovanie*, 2021, no. 4(79), pp. 40–44.

[6] Kotkova E.A., Matveev A.V. Metodika intellektual'nogo prognozirovaniya effektivnosti upravleniya evakuatsiey lyudey iz obshchestvennykh zdaniy [Methodology for intellectual forecastingof the efficiency of managing people evacuation of from public buildings]. *Vestnik Sankt-Peterburgskogo universiteta Gosudarstvennoy protivopozharnoy sluzhby MChS Rossii*, 2021, no. 4, pp. 107–120.

[7] Tsvirkun A.D., Rezchikov A.F., Samartsev A.A., et al. Integrirovannaya model' dinamiki rasprostraneniya opasnykh faktorov pozhara v pomeshcheniyakh i evakuatsii iz nikh [Integrated model of the fire dangerous factors dynamics in premises and the evacuation]. *Vestnik komp'yuternykh i informatsionnykh tekhnologiy*, 2019, vol. 2, no. 176, pp. 47–54.

[8] Tsvirkun A.D., Rezchikov A.F., Samartsev A.A., et al. Sistema integrirovannogo modelirovaniya rasprostraneniya opasnykh faktorov pozhara i evakuatsii lyudey iz pomeshcheniy [System of integrated simulation of spread of hazardous factors of fire and evacuation of people from indoors]. *Automation and Remote Control*, 2022, no 5, pp. 26–42.

[9] Samartsev A., Ivaschenko V., Rezchikov A., et al. Multiagent model of people evacuation from premises while emergency. *Advances in Systems Science and Applications*, 2019, vol. 19, no. 1, pp. 98–115.

[10] Gamayunova V.O., Bogomolov A.S., Kushnikov V.A., et al. Multiagentnoe modelirovanie evakuatsii iz pomeshcheniy s uchetom stolknoveniy agentov [Multi-agent modeling of evacuation from premises with consideration of agent collisions]. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*, 2025, vol. 25, no 1, pp. 106–115.

[11] Rosa A.C., Falqueiro M.C., Bonacin R., et al. EvacuAI: An Analysis of Escape Routes in Indoor Environments with the Aid of Reinforcement Learning. *Sensors*, 2023, vol. 23, no. 21, art. no. 8892.

[12] Ünal A.E., Gezer C., Pak B.K., et al. Generating emergency evacuation route directions based on crowd simulations with reinforcement learning. *2022 Innovations in Intelligent Systems and Applications Conference ASYU*. Antalya: IEEE, 2022, pp. 1–6.

[13] Xu D., Huang X., Mango J., et al. Simulating multi-exit evacuation using deep reinforcement learning. *Transactions in GIS*, 2021, vol. 25, no. 3, pp. 1542–1564.

[14] Malebary S.J., Basori A.H., Soliman alkayal E. Reinforcement learning for Pedestrian evacuation Simulation and Optimization during Pandemic and Panic situation. *Journal of Physics: Conference Series*, 2021, vol. 1817, no. 1, art. no. 012008.

[15] Komatsu H. Multi-agent reinforcement learning using echo-state network and its application to pedestrian dynamics. *arXiv preprint arXiv:2312.11834*, 2023.

[16] Sinpan N., Sasithong P., Chaudhary S., et al. Simulative Investigations of Crowd Evacuation by Incorporating Reinforcement Learning Scheme. *ICACS '22: Proceedings of the 6th International Conference on Algorithms, Computing and Systems.* New York: Association for Computing Machinery, 2022, pp. 1–5.

[17] Hassanpour S., Rassafi A.A., González V.A., et al. A hierarchical agent-based approach to simulate a dynamic decision-making process of evacuees using reinforcement learning. *Journal of choice modelling*, 2021, vol. 39, art. no. 100288.

[18] Schulman J., Wolski F., Dhariwal P., et al. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[19] Yu C., Velu A., Vinitsky E., et al. The Surprising Effectiveness of PPO in Cooperative Multi-Agent Games. *arXiv preprint arXiv:2103.01955*, 2022.

[20] Liu Z., Yao C., Na W., et al. MAPPO-Based Optimal Reciprocal Collision Avoidance for Autonomous Mobile Robots in Crowds. *2023 IEEE International Conference on Systems, Man, and Cybernetics SMC*. Honolulu: IEEE, 2023, pp. 3907–3912.
[21] Guo Y., Liu J., Yu R., et al. MAPPO-PIS: A Multi-agent Proximal Policy Optimization Method with Prior Intent Sharing for CAVs Cooperative Decision-Making. *Computer Vision – ECCV 2024 Workshops*. Cham: Springer Nature Switzerland, 2025, pp. 244–263.
[22] Shixin Z., Feng P., Anni J., et al. The unmanned vehicle on-ramp merging model based on AM-MAPPO algorithm. *Scientific Reports*, 2024, vol. 14, no. 1, art. no. 19416.
[23] Lowe R., Wu Y., Tamar A., et al. Multi-agent actor-critic for mixed cooperative-competitive environments. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook: Curran Associates Inc, 2017, pp. 6382–6393.
[24] Xiong J., Wang Q., Yang Z., et al. Parametrized Deep Q-Networks Learning: Reinforcement Learning with Discrete-Continuous Hybrid Action Space. *arXiv preprint arXiv:1810.06394*, 2018.
[25] Srivastava N., Hinton G., Krizhevsky A., et al. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 2014, vol. 15, no. 56, pp. 1929–1958.
[26] Gal Y., Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *Proceedings of The 33rd International Conference on Machine Learning*. New York: PMLR, 2016, pp. 1050–1059.
[27] Narvekar S., Peng B., Leonetti M., et al. Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey. *Journal of Machine Learning Research*, 2020, vol. 21, no. 181, pp. 1–50.
[28] Sutton R.S., Barto *A. Reinforcement learning: an introduction. Second edition.* Cambridge, The MIT Press, 2018, 552 p.
[29] Lillicrap T.P., Hunt J.J., Pritzel A., et al. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.
[30] Ob utverzhdenii svoda pravil SP 1.13130 «Sistemy protivopozharnoy zashchity. Evakuatsionnye puti i vykhody»: Prikaz MChS Rossii ot 19 marta 2020 g. N 194. Moskva. 2020.
[31] Trivedi A., Rao S. Agent-Based Modeling of Emergency Evacuations Considering Human Panic Behavior. *IEEE Transactions on Computational Social Systems*, 2018, vol. 5, no. 1, pp. 277–288.
[32] Ding N., Sun C. Experimental study of leader-and-follower behaviours during emergency evacuation. *Fire Safety Journal*, 2020, vol. 117, art. no. 103189.
[33] Wang L., Zheng J., Zhang X., et al. Pedestrians behavior in emergency evacuation: Modeling and simulation. *Chinese Physics B*, 2016, vol. 25, no. 11, art. no. 118901

**Silinskaya A.A.**, Master's student at Saratov State University. e-mail: asilinskaya@gmail.com

**Bogomolov A.S.,** Dr. Sci. (Eng.), Associate Professor, Professor at Saratov State University; Leading Researcher, Head of laboratory at Saratov Scientific Center of the Russian Academy of Sciences, e-mail: bogomolov@iptmuran.ru

**Kushnikov V.A.,** Dr. Sci. (Eng.), Professor, Professor at Saratov State University; Chief Researcher, Head of Laboratory at Saratov Scientific Center of the Russian Academy of Sciences; e-mail: kushnikoff@iptmuran.ru